Eigenvectors and diagonalization

- eigenvectors
- diagonalization
**Eigenvectors and eigenvalues**

λ ∈ ℂ is called an *eigenvalue* of \( A \in \mathbb{C}^{n \times n} \) if

\[
\chi(\lambda) = \det(\lambda I - A) = 0
\]

equivalent to:

- there exists nonzero \( v \in \mathbb{C}^n \) s.t. \((\lambda I - A)v = 0\), *i.e.*, \( Av = \lambda v\)

  any such \( v \) is called an *eigenvector* of \( A \) (associated with eigenvalue \( \lambda \))

- there exists nonzero \( w \in \mathbb{C}^n \) s.t. \( w^T(\lambda I - A) = 0\), *i.e.*, \( w^T A = \lambda w^T\)

  any such \( w \) is called a *left eigenvector* of \( A \)
Complex eigenvalues and eigenvectors

\[ A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \Rightarrow \lambda I - A = \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} \Rightarrow \lambda^2 + 1 = 0 \]
\[ \lambda = \pm i \]

\[ \text{if } \psi \text{ is an eigenvector of } A \text{ with eigenvalue } \lambda, \text{ then so is } \alpha \psi, \text{ for any } \alpha \in \mathbb{C}, \alpha \neq 0 \]
\[ A\psi = \lambda \psi \]
\[ A(\alpha \psi) = \lambda(\alpha \psi) \]

\[ \text{even when } A \text{ is real, eigenvalue } \lambda \text{ and eigenvector } \psi \text{ can be complex} \]

\[ \text{when } A \text{ and } \lambda \text{ are real, we can always find a real eigenvector } \psi \text{ associated with } \lambda: \text{ if } A\psi = \lambda \psi, \text{ with } \]
\[ A \in \mathbb{R}^{n \times n}, \lambda \in \mathbb{R}, \text{ and } \psi \in \mathbb{C}^n, \text{ then } \]
\[ A\psi = \lambda \psi \quad \text{so } \Re \psi \text{ and } \Im \psi \text{ are real eigenvectors, if they are nonzero} \]
\[ (\text{and at least one is}) \]

\[ \text{conjugate symmetry: if } A \text{ is real and } \psi \in \mathbb{C}^n \text{ is an eigenvector associated with } \lambda \in \mathbb{C}, \text{ then } \overline{\psi} \text{ is an} \]
\[ \text{eigenvector associated with } \overline{\lambda}: \]
\[ \text{taking conjugate of } A\psi = \lambda \psi \text{ we get } \overline{A\psi} = \overline{\lambda \psi}, \text{ so } \]
\[ A\psi = \lambda \psi \quad \Rightarrow \quad \overline{A\psi} = \overline{\lambda \psi} \quad \Rightarrow \quad \overline{A} \overline{\psi} = \overline{\lambda \psi} \quad \Rightarrow \quad A\overline{\psi} = \overline{\lambda \psi} \quad \Rightarrow \quad A\overline{\psi} = \overline{\lambda \psi} \]
\[ \text{we'll assume } A \text{ is real from now on . . .} \]
Scaling interpretation

\[
y = A\lambda
\]

\[
v \rightarrow y = Av = \lambda v
\]

(assume \(\lambda \in \mathbb{R}\) for now; we’ll consider \(\lambda \in \mathbb{C}\) later)

if \(v\) is an eigenvector, effect of \(A\) on \(v\) is very simple: scaling by \(\lambda\)

- \(\lambda \in \mathbb{R}, \lambda > 0\): \(v\) and \(Av\) point in same direction
- \(\lambda \in \mathbb{R}, \lambda < 0\): \(v\) and \(Av\) point in opposite directions
- \(\lambda \in \mathbb{R}, |\lambda| < 1\): \(Av\) smaller than \(v\)
- \(\lambda \in \mathbb{R}, |\lambda| > 1\): \(Av\) larger than \(v\)

(we’ll see later how this relates to stability of continuous- and discrete-time systems. . . )
Diagonalization

**NOT always possible**

suppose $v_1, \ldots, v_n$ is a *linearly independent* set of eigenvectors of $A \in \mathbb{R}^{n \times n}$:

$$Av_i = \lambda_i v_i, \quad i = 1, \ldots, n$$

express as

$$A \underbrace{[v_1 \quad \cdots \quad v_n]}_{\text{Suppose}} = \underbrace{[v_1 \quad \cdots \quad v_n]}_{\text{Suppose}} \underbrace{[\lambda_1 \quad \cdots \quad \lambda_n]}_{\text{Suppose}}$$

define $T = [v_1 \quad \cdots \quad v_n]$ and $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$, so

$$AT = T\Lambda$$

$$A = T \Lambda T^{-1}$$
Diagonalization

\[ A T = \Sigma \Lambda \]

\[ T^{-1} A T = \Lambda \]

- \( T \) invertible means \( v_1, \ldots, v_n \) linearly independent
- similarity transformation by \( T \) diagonalizes \( A \)
- existence of invertible \( T \) such that
  \[ T^{-1} A T = \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \]
  is equivalent to existence of a linearly independent set of \( n \) eigenvectors
- we say \( A \) is diagonalizable
- if \( A \) is not diagonalizable, it is sometimes called defective

\[ A \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ 0 & \cdots & \lambda_n \end{bmatrix} \]

\[ A v_i = \lambda_i v_i \]

\( i = 1, \ldots, n \)
Not all matrices are diagonalizable

example: \[ A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \]

- characteristic polynomial is \[ \chi(s) = s^2 \], so \( \lambda = 0 \) is only eigenvalue
- eigenvectors satisfy \( Au = 0v = 0 \), i.e.
  \[
  \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0
  \]
- so all eigenvectors have form \( v = \begin{bmatrix} v_1 \\ 0 \end{bmatrix} \) where \( v_1 \neq 0 \)
- thus, \( A \) cannot have two independent eigenvectors
Distinct eigenvalues

if $A$ has distinct eigenvalues then $A$ is diagonalizable

- distinct eigenvalues means $\lambda_i \neq \lambda_j$ for $i \neq j$
- the converse is false — $A$ can have repeated eigenvalues but still be diagonalizable

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A \mathbf{v} = I \mathbf{v} = \lambda \mathbf{v} = \mathbf{v}$$

$$I \mathbf{v} = \mathbf{v}$$

$$T = I \rightarrow \Lambda = T^{-1} A T$$
Diagonalization and left eigenvectors

Rewrite $T^{-1}A T = \Lambda$ as $T^{-1} A = \Lambda T^{-1}$, or

$A T = T \Lambda$

where $w_1^T, \ldots, w_n^T$ are the rows of $T^{-1}$

Thus

$w_i^T A = \lambda_i w_i^T$

i.e., the rows of $T^{-1}$ are (lin. indep.) left eigenvectors, normalized so that

$w_i^T v_j = \delta_{ij}$

($i.e.$, left & right eigenvectors chosen this way are dual bases)

$$I = \begin{bmatrix} w_1^T \\ \vdots \\ w_n^T \end{bmatrix} \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix} = \begin{bmatrix} w_1^T v_1 & \ldots & w_1^T v_n \\ \vdots & \ddots & \vdots \\ w_n^T v_1 & \ldots & w_n^T v_n \end{bmatrix}.$$